Theor Appl Genet (2004) 108:423–432
DOI 10.1007/s00122-003-1455-1

ORIGINAL PAPER

W. S. De Jong · N. T. Eannetta · D. M. De Jong · M. Bodis

Candidate gene analysis of anthocyanin pigmentation loci in the *Solanaceae*

Received: 24 May 2003 / Accepted: 28 August 2003 / Published online: 2 October 2003

Abstract Crop species in the *Solanaceae*, which includes tomato (*Lycopersicon esculentum*), potato (*Solanum tuberosum*), pepper (*Capsicum spp.*), and eggplant (*S. melongena*), exhibit natural variation in the types, levels, and tissue-specific expression patterns of anthocyanin pigments. While the identities of the genes underpinning natural variation in anthocyanin traits in these crops are largely unknown, many structural genes and regulators of anthocyanin biosynthesis have been isolated from the solanaceous ornamental species *Petunia*. To identify candidate genes that may correspond to loci controlling natural variation in the four crops, 13 anthocyanin-related genes were localized on a tomato F_2 genetic map. Gene map positions were then compared to mapped mutants in tomato and through comparative genetic maps to natural variants in potato, eggplant, and pepper. Similar map positions suggest that the tomato mutants *anthocyaninless*, *entirely anthocyaninless*, and *anthocyanin gainer* correspond to flavonoid 3′5′-hydroxylase (*f3′5′h*), anthocyanidin synthase, and the *Petunia* Myb domain trancriptional regulatory gene *an2*, respectively. Similarly potato *R*, required for the production of red pelargonidin-based pigments, *P*, required for production of purple delphinidin-based pigments, and *I*, required for tissue-specific expression in tuber skin, appear to correspond to dihydroflavonol 4-reductase, *f3′5′h* and *an2*, respectively. The map location of *an2* also overlaps pepper *A* and eggplant *fap*10.1, *lla*10.1, *lra*10.1, *sa*10.1, *pa*10.1 and *ca*10.1, suggesting that a homologous regulatory locus has been subjected to parallel selection in the domestication of many solanaceous crops. To test the hypothesis that tomato *anthocyaninless* corresponds to *f3′5′h*, a portion of the gene was sequenced. A premature stop codon was observed in an *anthocyaninless* mutant, but not in wild-type.

Introduction

Red and purple anthocyanin pigments are important constituents of many ornamental and agricultural crops. In addition to frequently providing the basis for visually appealing colors in flowers, fruits, leaves, stems, tubers, and roots, anthocyanin pigments may also act as valuable antioxidants in the human diet (Noda et al. 2000). While the genetics of anthocyanin synthesis and its regulation are well-characterized in several model systems like *Petunia* and maize (Holton and Cornish 1995), in most crops the identity of the genes underpinning natural variation in the types, levels, and tissue-specific expression patterns of anthocyanin pigments are still not known.

Over a dozen genes that influence anthocyanin pigmentation have been isolated from *Petunia*. Cloned *Petunia* genes include those encoding most of the enzymes of the biosynthetic pathway, e.g., dihydroflavonol 4-reductase (Beld et al. 1999) and flavonoid 3′, 5′-hydroxylase (Holton et al. 1993), as well as several transcriptional regulators, e.g., the Myb domain gene *an2* (Quattrocchio et al. 1999). As a member of the *Solanaceae*, *Petunia* is closely related to many important crop plants, including potato (*Solanum tuberosum*), tomato (*Lycopersicon esculentum*), pepper (*Capsicum spp.*), and eggplant (*S. melongena*). There is considerable variation for anthocyanin accumulation in each of these crop species and in some cases, this variation is economically important. The skin of potato tubers, for example, can accumulate either red, purple, or no anthocyanin pigments. The potato *R* locus is required for the production of red anthocyanins, which have been shown to be derivatives of pelargonidin, while *P* is required for the synthesis of purple pigments, which are typically derived from petunidin (Dodds and Long 1955; De Jong 1991; Lewis et al. 1998). A third locus, *I,* is required for the synthesis of red or purple anthocyanins in tuber skin

Communicated by H.F. Linskens

W. S. De Jong (✉) · N. T. Eannetta · D. M. De Jong · M. Bodis
Department of Plant Breeding,
Cornell University, Ithaca, NY 14853-1901, USA
e-mail: wsd2@cornell.edu
Tel.: +1-607-2545384
Fax: +1-607-2556683

(Dodds and Long 1956). All three of these loci have been mapped in the potato genome (Van Eck et al. 1993, 1994). Over a dozen loci affecting anthocyanin accumulation have been described in tomato (von Wettstein-Knowles 1967), and several have been mapped relative to molecular markers (Tanksley et al. 1992). Several classical loci have also been reported in pepper (Daskalov and Poulos 1994), and a number of quantitative trait loci have recently been described in eggplant (Doganlar et al. 2002b).

The genomes of tomato, potato, pepper, and eggplant have been shown, with the exception of several chromosomal rearrangements, to share extensive co-linearity of gene order (Tanksley et al. 1992; Livingstone et al. 1999; Doganlar et al. 2002a). Thus, once the position of a gene has been determined in any one of these species, it is possible to infer the location in the remaining three. In order to identify candidate genes that may correspond to classical loci influencing anthocyanin accumulation in these crop plants, we report here the genetic map locations in tomato of thirteen genes known to be involved in anthocyanin biosynthesis or its regulation in *Petunia*. Our primary motivation was to identify candidate genes for the *R*, *P*, and *I* loci in potato, and excellent positional candidates were identified for all three. Strong candidates were also found for several other mapped loci in tomato, pepper, and eggplant. In conjunction with a similar study that reported the mapping of carotenoid structural genes in pepper (Thorup et al. 2000), this work illustrates the power that mapping genes underlying a defined biochemical pathway can bring to the identification of genes responsible for natural variation along that pathway, both in the species where the mapping is performed, as well as in related species that are connected to it through comparative genetic maps.

Materials and methods

Genetic mapping of anthocyanin-related loci

Gene map locations were determined by following RFLP segregation in a tomato F_2 mapping population of 83 individuals. The parents of the population were *Lycopersicon esculentum* LA925 and *L. pennellii* LA716 (Fulton et al. 2002). Prior to mapping, each probe was first hybridized to a survey blot containing genomic DNA of the two parents digested separately with *Eco*RI, *Eco*RV, *Dra*I, *Hae*III, and *Sca*I. The probe was then hybridized to DNA blots where genomic DNA of the 83 progeny had been digested with the restriction enzyme that revealed the clearest polymorphism between the two parents. Random hexamer labeling of probes, hybridization, and washing methods were as described previously (Feinberg and Vogelstein 1983). Filters were washed to a final stringency of 0.5× SSC (saline sodium citrate) at 65°C (1× SSC is 0.15 M NaCl and 0.015 M trisodium citrate). Gene map locations relative to a large number of previously mapped RFLP markers (Fulton et al. 2002; http://www.sgn.cornell.edu) were determined using Mapmaker software (Lander et al. 1987).

Four loci were mapped using existing tomato cDNA clones with clear sequence similarity to known *Petunia* genes. Plasmid cTOB11G7 was used as a probe to map chalcone isomerase (*chi*), cLED38K19 was used to locate flavonol synthase (*fls*), cLEB1D3 was used to map anthocyanidin 3-O-glucosyltransferase (*3gt*), and cLED11P23 was used to define a tomato locus related to the *jaf13* gene of *Petunia*. All four plasmids were kindly provided by S.D. Tanksley (Cornell University).

Eight loci corresponding to chalcone synthase (*chs*), dihydroflavonol 4-reductase (*dfr*), flavonoid 3′, 5′-hydroxylase (*f3′5′h*), anthocyanidin synthase (*ans;* also known as leucoanthocyanidin dioxygenase), anthocyanidin-3-glucoside rhamnosyltransferase (*rt*), *an1*, *an2*, and *an11* were mapped using probes generated by PCR amplification of known *Petunia*, tomato, or potato genes (Table 1). The flavanone 3-hydroxylase (*f3 h*) probe was generated by PCR amplification of tomato genomic DNA using primers designed against two tomato ESTs with >93% sequence identity to the known potato *f3 h* gene (AY102035) (Table 1). The *an2* probe was obtained by PCR amplification of the *Petunia* gene using primers *an2*-5′ and *an2*-3′ as described in Quattrocchio et al. (1998). With the exception of *an1* and *f3 h*, all amplified gene fragments were ligated into cloning vector pCR2.1 (Invitrogen). The *an1* gene fragment was ligated into vector pGEM-T (Promega). Each cloned gene fragment was partially sequenced to confirm gene identity

Table 1 Primers used for amplification of anthocyanin gene probes

Gene	GenBank accession no.	Species	Fragment size (bp)	Primer sequence (5′–3′)
chs	U47738	Potato	≈600	F GCGACTCCTTCGAACTGTG R AAGTTTTTCGGGCTTTAGGC
f3h	BG132926 (F) BG735030 (R)	Tomato	≈1800	F GGGCACCTTCAACACTA R ATCGGCTCATCCATTAT
dfr	Z18277	Tomato	≈1500	F CACTCTCCTCCGAAGACGAC R TCCATTGTCTGCAGTGCTTC
f3′5′h	AF081575	Petunia	≈600	F TTTGTTCACAGCTGGTACGG R AGAGGGACAGCTTTCTGCAA
ans	X70786	Petunia	≈1000	F AAGGAGATTCGCGAGAAATG R GCCACACTGTTCATCCTCCT
rt	X71060	Petunia	≈1500	F GCTGAGCTTCTCAAGGTTGC R ACCATCACCTTTTCCACAGC
an11	U94748	Petunia	≈1000	F TCAATCCCACCCACCACTAT R CCCATCATCTCCACCTGAAC
an1	AF260918	Petunia	≈1300	F CGGCCCTAGTTATGATGAATTATC R ACCTCCACTTTAAGTTCCCTTAGC

prior to genetic mapping. Uncloned *f3 h* PCR products were also sequenced prior to mapping. For *chs*, the observed amplification product, approximately 600 bp in length, was shorter than expected. The expected amplification product size was approximately 1,000 bp, plus any introns that might be present in the amplified genomic sequence. Because sequencing of the cloned 600-bp product revealed several tracts with 100% sequence identity to U47738, this amplification product was nevertheless deemed acceptable for mapping purposes.

Tomato and potato sequences similar to known anthocyanin genes were identified by searching GenBank sequences during May 2003 with the program BLAST (http://www.ncbi.nlm.nih.gov/blast/). An EST sequence was considered to be similar to the query if it returned an expect score of less than 10^{-8}. At this time there were approximately 155,000 tomato and 90,000 potato EST sequences present in GenBank.

A fragment of the tomato *f3′5′h* gene was amplified with forward primer f3′5′h-i (GTG GCC GGT GAT CGG AGC ACT) and reverse primer f3′5′h-d (AAG CCT TTC TCC TTC AGA ATT). These primers were designed against *f3′5′h* sequences conserved in *Petunia*, eggplant, and *Lycianthes*, and generated a product approximately 1.6 kb in length, representing approximately 1/2 of the predicted coding sequence of the tomato gene. When uncloned PCR products were directly sequenced with f3′5′h-i and f3′5′h-d primers, sequence corresponding to about 1/3 of the tomato gene was obtained. Tomato monogenic mutant accessions 3–073, 3–414, LA1048 and LA1049 were obtained from the Tomato Genetics Resource Center (University of California, Davis).

Results

Map locations

A schematic diagram of the initial stages of the anthocyanin biosynthetic pathway is shown in Fig. 1. Genetic map locations for thirteen anthocyanin-related genes in tomato are illustrated in Fig. 2. Additional details relating to the mapping of each gene follow below.

Chalcone synthase

The probe for mapping *chs* was obtained by PCR amplification of a known potato gene (GenBank accession U47738). The hybridization data were consistent with two unlinked loci segregating in the population, but only one of the two gave a hybridization signal sufficiently strong to be unambiguously scored. Two *chs* loci have previously been mapped in tomato by O'Neill et al. (1990); TCHS1 mapped to chromosome 9, while TCHS2 mapped to the lower arm of chromosome 5. Tanksley et al. (1992) reported the location of three tomato *chs* genes. CHS1 mapped to the top of chromosome 6, CHS3 mapped to the lower arm of chromosome 5, and CHS4 mapped to the lower arm of chromosome 9. Given similar map locations the *chs* locus mapped in the present study to a 7.1 cM interval between TG23 and TG60 is likely to correspond to TCHS2 of O'Neill et al. (1990) and CHS3 of Tanksley et al. (1992). Examination of current (May 2003) tomato EST sequences in GenBank revealed a third sequence (accession AW979450) with approximately 80–85% identity to TCHS1 and TCHS2. A potato *chs*-like gene distinct from TCHS1, TCHS2 and AW979450 can also be found in GenBank (accession BO114342). At present, 11 of 15 tomato ESTs with clear DNA sequence similarity to *Petunia chsA* (accession X14591) correspond to TCHS2, suggesting that TCHS2 is the most actively transcribed *chs* locus in tomato.

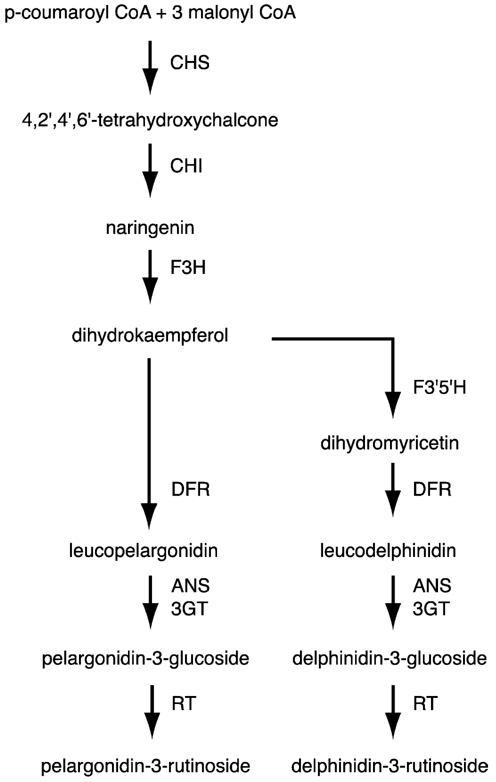

Fig. 1 Schematic diagram illustrating the initial stages of the anthocyanin biosynthetic pathway. Enzyme abbreviations are as follows: *CHS* chalcone synthase, *CHI* chalcone isomerase, *F3H* flavanone 3-hydroxylase, *DFR* dihydroflavonol 4-reductase, *F3′5′H* flavonoid 3′, 5′-hydroxylase, *ANS* anthocyanidin synthase, *3GT* UDP-glucose anthocyanidin 3-O-glucosyltransferase, *RT* UDP-rhamnose anthocyanidin-3-glucoside rhamnosyltransferase

Chalcone isomerase

The partially sequenced probe (accession AW623561) used to map *chi* in tomato shares approximately 80% sequence identity with *Petunia chiA* (accession X14589) and 87% sequence identity with *chiB* (accession X14590). Similar to the *chs* probe, the *chi* probe revealed two unlinked loci in tomato, only one of which yielded a sufficiently strong hybridization signal for mapping. The mapped *chi* resides on chromosome 5, 0.7 cM south of CT93. Searching GenBank ESTs with AW623561 as the query revealed at least two *chi*-like genes in tomato. Thirteen of 16 tomato *chi*-like ESTs appear to be transcribed from the same locus as AW623561 since they share 95% or greater sequence identity with it. The other three ESTs are derived from a distinct gene typified by accession BI930695, which shares 85% sequence identity with AW623561.

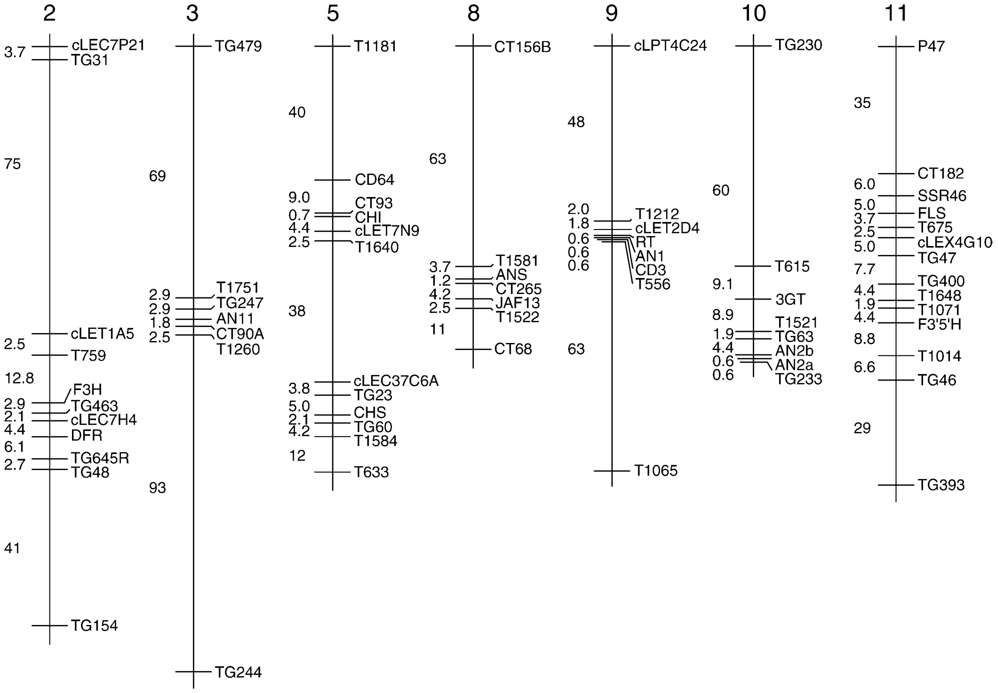

Fig. 2 Genetic map locations of anthocyanin-related structural and regulatory genes relative to previously mapped framework RFLP markers. Each chromosome, *numbered* at top, shows the framework markers most tightly linked to the mapped anthocyanin pathway genes, as well as the most distal framework markers. Genetic distances (in cM) between loci are shown to the *left* of each chromosome. All framework markers were mapped with LOD scores >3. The following genes were also mapped with LOD>3: *an1*, *ans*, *jaf13*, *3gt*. These genes were mapped with 2<LOD<3: *an2*a, *an2*b, *an11*, *chs*, *f3 h*, *fls*. The remaining genes were mapped with LOD<2: *chi*, *dfr*, *rt*, *f3′5′h*

Flavonol synthase

This gene is not required for production of anthocyanin pigments. It was mapped because it competes with the gene products of *f3′5′h* and *dfr* for common dihydroflavonol substrates and thus might influence anthocyanin synthesis. A partially sequenced tomato cDNA clone (accession AI771853) that shares approximately 85% sequence identity with *Petunia fls* (accession Z22543) was used as a mapping probe. The RFLP banding pattern was simple and suggests that tomato *fls* is a single copy gene, located on chromosome 11. The single copy nature of *fls* is further supported by the observation that all 25 tomato ESTs currently in GenBank with nucleic acid sequence similarity to *Petunia fls* share 97–100% sequence identity with each other.

Flavanone 3-hydroxylase

A PCR product from tomato, sharing >90% sequence identity with the known potato *f3h* gene (AY102035), was used for mapping. The simple banding patterns observed suggest that tomato encodes a single *f3h* gene, located on chromosome 2. The 15 tomato ESTs currently in Genbank with similarity to potato *f3h* share 97–100% sequence identity with each other and are thus presumably transcribed from a single gene.

Dihydroflavonol 4-reductase

The *dfr* gene has previously been mapped to a 37.8 cM interval on chromosome 2 by Bongue-Bartelsman et al. (1994). Using tomato *dfr* as a probe, we confirmed the chromosome 2 localization, and by virtue of using a higher-resolution reference mapping population, have narrowed the map location to a smaller 10.5 cM interval between cLEC7H4 and TG645R, approximately 9 cM south of *f3 h.* Similar to Bongue-Bartelsman et al. (1994), we observed a simple hybridization pattern, indicating that *dfr* is present in only one copy in the tomato genome. No tomato ESTs corresponding to DFR were present in GenBank as of May 2003, although two ESTs from potato have been deposited (accessions BG595686 and BE923282).

Flavonoid 3′,5′-hydroxylase

A fragment of the *Petunia Hf1* gene was used as a probe to map the corresponding locus in tomato. *Hf1* encodes a cytochrome P450 monooxygenase gene (Holton et al. 1993) and revealed a complex banding pattern during RFLP analysis, suggesting the presence of numerous related genes. The high copy number was not unexpected since approximately 286 cytochrome P450 genes are known to be present in *Arabidopsis* (The Arabidopsis Genome Initiative 2000). Despite the high copy number, one or two bands hybridized much more strongly than all others on the survey blot. The most strongly hybridizing band was polymorphic following digestion with *Dra*I and was readily mapped to the lower arm of chromosome 11. No tomato or potato ESTs with extensive sequence similarity to *Petunia Hf1* have yet been described.

Anthocyanidin synthase

A fragment of the *Petunia ans* gene (accession X70786) was used to map the related locus in tomato. The RFLP banding pattern was simple and suggests that tomato encodes a single *ans* locus, which maps to the lower arm of chromosome 8. No tomato or potato ESTs with detectable nucleotide sequence similarity to *Petunia ans* have yet been described.

UDP-glucose anthocyanidin 3-O-glucosyltransferase

A partially sequenced tomato cDNA clone (accession AI482628) that shares approximately 81% sequence identity with *Petunia 3gt* (accession AB027454) was used as a mapping probe. Tomato *3gt* mapped to the lower arm of chromosome 10, and based on the simple RFLP banding patterns observed, is likely a single copy gene. There are currently 20 tomato ESTs in GenBank with high levels of sequence similarity to *Petunia 3gt*. Most of these ESTs cluster into a contig corresponding to the 5′ half of *Petunia 3gt* while the remainder cluster into a contig corresponding to the 3′ end. Since the two contigs do not overlap in sequence it is not yet possible to conclude that they are derived from the same gene.

UDP-rhamnose anthocyanidin-3-glucoside rhamnosyltransferase

A cloned fragment of the *Petunia rt* gene (X71060) was used as a probe to map the corresponding locus in tomato. Tomato *rt* maps to the middle of chromosome 9. Based on the simple RFLP banding patterns observed, *rt* is likely encoded by a single locus. Of the 16 tomato ESTs currently in GenBank with similarity to *Petunia rt*, 14 can be placed into a single contig and are presumably derived from a single gene. The remaining two ESTs (BM410878 and BM535249) contain internal polyA tracts and appear to represent chimeric cDNA clones.

An1

A fragment of the *Petunia an1* gene, a basic helix-loop-helix (bHLH) transcriptional regulator of *dfr* (Spelt et al. 2000) and possible regulator of other structural genes as well, was used to determine the map location of the corresponding tomato locus. Tomato *an1* mapped to chromosome 9, 0.6 cM south of *rt*. The simple RFLP banding pattern suggested that tomato encodes a single *an1*-like gene. No tomato or potato ESTs with extensive DNA sequence similarity to *an1* have yet been deposited in GenBank.

Jaf13

The *jaf13* gene encodes a second bHLH regulator of anthocyanin gene expression in *Petunia* (Quattrocchio et al. 1998). Using a partially sequenced tomato cDNA clone (accession AI486789) that shares 87% sequence identity with *jaf13* as a probe revealed a single locus in tomato, near the lower tip of chromosome 8 and slightly south of *ans*. The 11 current tomato ESTs with high levels of sequence similarity to *jaf13* match three non-overlapping regions of this gene. While they are likely to be transcribed from a single locus, they are not contiguous so it is not yet possible to conclude this with certainty.

An2

The *Petunia an2* gene encodes a Myb domain transcriptional regulator of the anthocyanin pathway (Quattrocchio et al. 1999). An *an2* probe revealed two tightly linked tomato loci on the lower arm of chromosome 10. These two loci were separated by a single recombination event in the 83 progeny of the mapping population. Three tomato ESTs with similarity to *an2* have been deposited in GenBank. Typified by accession BE462282, these three share 97–99% sequence identity with each other.

An11

A fragment of *an11*, a WD40 repeat regulator of anthocyanin expression in *Petunia* (de Vetten et al. 1997), was used to map a related locus in tomato. Tomato *an11* maps to chromosome 3. Although it could not be mapped, an additional weakly hybridizing suggested that tomato might encode a second *an11*-like locus. Consistent with this possibility the ten current tomato ESTs with sequence similarity to *an11* can be divided into two sequence classes. Nine of them, exemplified by accession BI421668, appear to be derived from a single gene. Accession AW932616 shares only 90% sequence identity with these eight and thus must be transcribed from a distinct gene.

Associations with mapped classical and QTL loci

Many classical and quantitative trait loci influencing anthocyanin accumulation have previously been mapped in the tomato, potato, pepper, and eggplant genomes. In several cases, the genes mapped in this study are located sufficiently close to known functional loci to suggest the identity of the genes that these loci encode. Instances where genes mapped close to known anthocyanin-related loci are summarized in Table 2. These putative associations are detailed further in the Discussion.

Two of the associations in tomato—between the *anthocyanin reduced* locus and *f3h*, and between the *anthocyaninless* locus and *f3′5′h*—were tested further by

Table 2 Putative associations between mapped genes and mapped classical or quantitative trait loci influencing anthocyanin accumulation in four solanaceous crops[a]

Gene	Tomato[b]	Potato	Pepper	Eggplant
Chalcone synthase	-	-	-	-
Chalcone isomerase	*af* [c] ?	-	-	-
Flavonol synthase	-	-	-	-
Flavanone 3-hydroxylase	*are* [d]	-	-	-
Flavonoid 3′5′-hydroxylase	*a* [c]	*P* [f]	-	-
Dihydroflavonol 4-reductase	*aw* [e]	*R* [f]	-	-
Anthocyanidin synthase	*ae* [c]	-	-	-
UDP-glucose: anthocyanidin 3-O-Glucosyltransferase	-	-	-	-
UDP-rhamnose: anthocyanidin-3-Glucoside rhamnosyltransferase	-	-	-	-
an1	-	-	-	-
an2	*ag* [c]	*F* [f],*I* [g]	*A* [h]	*fap*10.1[i], *lla*10.1[i], *lra*10.1[i], *sa*10.1[i], *pa*10.1[i], *ca*10.1[i]
an11	-	-	-	-
jaf13	-	-	-	-

[a] Many *Petunia* anthocyanin mutants and the genes they correspond to are summarized in Holton and Cornish (1995).

[b] Tomato classical mutant abbreviations: *af anthocyanin free*, *are anthocyanin reduced*, *a anthocyaninless* , *aw anthocyanin without* , *ae entirely anthocyaninless* , *ag anthocyanin gainer*

[c] Mapped by Tanksley et al. (1992)

[d] Mapped by Borgnino et al. (1973)

[e] Demonstrated through complementation by Goldsbrough et al. (1994)

[f] Mapped by Van Eck et al. (1993)

[g] Mapped by Van Eck et al. (1994)

[h] Mapped by Ben Chaim et al. (2003)

[i] Mapped by Doganlar et al. (2002b)

comparing the sequence of *f3h* and *f3′5′h* in wild-type and mutant backgrounds. The *f3h* primers listed in Table 1 were used to amplify and sequence a fragment predicted to contain about 3/4 of the gene from tomato accessions 3–073, 3–414, and LA1049. By way of reference, the potato *f3h* gene (AY102035) encodes a protein 358 amino acids in length. Accessions 3–073 and 3–414 carry independent EMS-induced mutations at the *anthocyanin reduced* and *anthocyaninless* loci, respectively, and are both in a genetic background of VF36. LA1049 carries a radiation-induced mutation at the *anthocyanin free* locus in a Red Cherry background. Both 3–414 and LA1049 were assumed to be wild-type for *f3h*, since the mutations affecting anthocyanin production in these two accessions are present on different chromosomes than *anthocyanin reduced*. Direct sequencing of uncloned PCR products from the three accessions yielded sequence corresponding to codons 20 through 297 of potato *f3h* and revealed that 3–073 differed at a single nucleotide position from both 3–414 and LA1049 (Fig. 3). This G to A mutation is predicted to alter the amino acid sequence, from serine to asparagine, at a position corresponding to residue 116 of potato F3H (Fig. 3). This serine is conserved in potato (AAM48289), petunia (AAC49929), maize (AAA91227) and *Arabidopsis* (AAC68584), as well as in many other plants, so that the mutation likely interferes with F3H function. Indeed, Yoder et al. (1994) have previously reported that F3H enzyme activity is abolished in an *anthocyanin reduced* mutant.

flavanone 3-hydroxylase

```
wt      AAG AAA GGT GGC TTC ATT GTC TCA AGC CAC TTA CAG
3-073   AAG AAA GGT GGC TTC ATT GTC TCA AAC CAC TTA CAG
         K   K   G   G   F   I   V   S  S>N  H   L   Q
```

flavonoid 3’,5’-hydroxylase

```
wt      TCG ATG TTC GAT GCA AGC CAG GAT GGC GAA TGT GTG
3-414   TCG ATG TTC GAT GCA AGC TAG GAT GGC GAA TGT GTG
         S   M   F   D   A   S  O>*  D   G   E   C   V
```

Fig. 3 Mutations observed in the flavanone 3-hydroxylase gene of an *anthocyanin reduced* mutant and in the flavonoid 3′,5′-hydroxylase gene of an *anthocyaninless* mutant. (*Top*) DNA sequence from tomato accessions with wild-type (3–414 and LA1049) and mutant (3–073) *f3h* genes. The region shown corresponds to amino acids 108–119 of the potato *f3h* gene product; potato numbering is used because the tomato gene has not been completely sequenced. The mutation present in 3–073 is predicted to change a codon for serine (AGC) into a codon for asparagine (AAC). (*Bottom*) DNA sequence from tomato accessions with wild-type (3–073) and mutant (3–414) *f3′5′h* genes. The region shown corresponds to amino acids 159–170 of the *Petunia f3′5″h* gene product; *Petunia* numbering is used because the tomato gene has not been completely sequenced. The mutation present in 3–414 is predicted to change a codon for glutamine (CAG) into a stop codon (TAG). Predicted amino acid sequences (*one letter code*) are shown *below* the aligned DNA sequences

To test whether tomato *anthocyaninless* has a lesion in *f3′5′h*, two primers (see “Materials and methods”) based on regions conserved at the DNA level between known *f3’5’h* genes in *Petunia*, eggplant, and *Lycianthes* were used to amplify a fragment corresponding to approximately 1/2 of the tomato gene from *anthocyaninless* mutants 3–414 and LA1048 (genetic background: Kokomo), as well as from *anthocyanin reduced* mutant 3–073. By way of reference, the *Petunia f3′5′h* gene product is 506 amino acids long. Direct sequencing of uncloned tomato PCR products in a region corresponding to codons 105 through 263 of *Petunia f3′5′h* revealed that the tomato gene shares 84% nucleotide and 88% amino acid identity with the *Petunia* gene and indicated that accession 3–414 differed at a single nucleotide position from both 3-073 and LA1048 (Fig. 3). The C to T mutation in 3–414 creates a premature stop codon relative to 3–073 and LA1048, at a position that corresponds to amino acid 165 of *Petunia* F3′5′H. The presence of this stop codon provides excellent support for the hypothesis that *anthocyaninless* corresponds to *f3′5′h*. No mutations were observed in LA1048 relative to 3–073. Nevertheless, because only one third of the gene was sequenced, the presence of mutations elsewhere in LA1048 *f3′5′h* cannot be excluded.

Discussion

We have reported here the genetic map locations in tomato of thirteen genes related to anthocyanin biosynthesis. Eight of these genes encode enzymes of the anthocyanin biosynthetic pathway (*chs*, *chi*, *f3h*, *dfr*, *f3′5′h*, *ans*, *3gt*, *rt*), four encode regulatory genes (*an1*, *an2*, *an11*, *jaf13*) and one codes for a enzyme (*fls*) which competes with *dfr* for dihydroflavonols, intermediate metabolites in anthocyanin synthesis.

Our primary purpose in mapping these genes was to identify candidates for the *R*, *P*, and *I* loci in potato, which are responsible for natural variation in potato skin color. We chose to perform the gene mapping in tomato rather than potato because a tomato mapping population was available to us when we initiated the project, while a potato population was not. The extensive conservation of marker order between tomato and potato (Tanksley et al. 1992) led us to expect that we would be able to readily interpret the results in potato. This was indeed the case, although one unexpected difficulty was encountered on chromosome 2 (see below). Given the genetic similarities between tomato and pepper and between tomato and eggplant, we were also able to make specific predictions about the genes that correspond to several other known anthocyanin loci in pepper and eggplant, as well as in tomato itself. This work has thus resulted in the development of many specific genetic hypotheses, e.g., that the *R* locus in potato encodes *dfr*, or that the tomato *ae* locus encodes *ans*. Naturally, additional experimentation will be required to test each of these hypotheses more rigorously.

Association between potato *R* and *dfr*

The *R* locus is required for the production of red pelargonidin-related anthocyanin pigments and has previously been mapped in potato to a 24 cM interval between markers TG20(b) and STF13 (Van Eck et al. 1993). Neither of these markers is shared with the tomato map used in this study. The only shared markers are TG31, 41 cM to the north of *R*, and TG48, 33 cM to the south (Van Eck 1995). In the present study, the tomato *dfr* gene was mapped 87 cM south of TG31 and 9 cM north of TG48. At first glance, the widely differing positions of *dfr* and *R* appear to rule out a relationship. However, we have recently shown (De Jong et al. in press) that *dfr* cosegregates with *R* in three diploid potato populations, which serves to highlight an inconsistency in the current tomato/potato comparative map. In two independent potato maps (Bonierbale et al. 1988; Van Eck 1995; Van Eck et al. 1995) TG48 has mapped at or very close to the southern telomere, while in tomato TG48 is located 40 cM north of the telomere (Tanksley et al. 1992). One possible explanation, not previously suggested in the literature, is that part of the lower arm of potato chromosome 2 is inverted relative to tomato. Such an inversion would account for the apparent differences in map location of potato and tomato *dfr*. On a different note, Van Eck et al. (1993) have previously reported that a *chs* locus cosegregates with *R*. Since *chs* acts at a very early stage of anthocyanin synthesis, before the pelargonidin, cyanidin, and petunidin biosynthetic pathways diverge, and *R* manifests a pelargonidin-specific phenotype, it is unlikely that *chs* corresponds to *R*.

Association between potato *P* and *f3′5′h*

The *P* locus is required for the production of petunidin-based anthocyanins and has previously been mapped 36 cM south of TG47 and 6 cM north of TG30 (Van Eck et al. 1993). In the current study, *f3′5′h* mapped 18.4 cM south of TG47, while TG30 was not mapped. Petunidin is differentiated from pelargonidin in part by the presence of two additional hydroxyl groups on the B phenyl ring. In *Petunia*, these differential hydroxylation reactions are catalyzed by the gene product of *f3′5′h* (Holton et al. 1993). The similar genetic map locations and biochemical precedent thus suggest that the dominant allele at the *P* locus encodes an enzyme with flavonoid 3′,5-hydroxylase activity, while the recessive *p* allele lacks this activity.

Association between potato *I/F and Petunia an2*

The *I* locus is required for the synthesis of anthocyanin pigments in tuber skin. Linkage analyses have shown *I* is tightly linked to several other loci that also mediate tissue-specific anthocyanin expression. The *F* locus, for example, is required for expression in potato flowers, while a series of alleles at the *B* locus conditions a hierarchical pattern of expression in the floral abscission layer, tuber eyebrow, cotyledonary node of the embryo, and the base of petioles (Dodds and Long 1956). *I* and *F* are separated by about 1.7 cM (Dodds and Long 1956). *F* has been mapped 3 cM north of TG63 (Van Eck et al.

1993). The previous mapping of *I* by Van Eck et al. (1994) shared no markers with the current study. Nevertheless, the purple skin color locus localized by Gebhardt et al. (1991) likely corresponds to *I* and mapped about 3 cM south of TG63. Tissue-specific anthocyanin expression in several model systems is known to result from the tissue-specific expression of transcriptional regulators, e.g., the Myb domain gene *an2*, required for anthocyanin expression in the floral limb, or the maize bHLH domain gene *B*, required for expression in the seed and some vegetative tissues (Selinger and Chandler 1999). Of the four *Petunia* regulatory genes mapped in tomato, only *an2* mapped close to the telomere of the lower arm of chromosome 10, 4.4 cM and 5.0 cM south of TG63. It is intriguing that the two tightly linked loci revealed by the *an2* probe mirrors the tight linkage observed between *F* and *I*. Given the precedent of *an2* controlling tissue-specific expression of anthocyanin expression in *Petunia*, and the mapping of *an2, F*, and *I* close to TG63 in the tomato and potato genomes, potato genes related to *an2* constitute good candidates for *F* and *I*.

Association between tomato *ag* and *Petunia an2*

The tomato *anthocyanin gainer* (*ag*) locus is characterized by delayed expression of anthocyanins during plant development. *Ag* has previously been mapped close to the telomere of the lower arm of tomato chromosome 10, cosegregating with RFLP marker TG233 (Tanksley et al. 1992). Since *an2* mapped very close to TG233 in the present study, either of the two tomato loci defined by the *an2* probe constitute excellent candidates for *ag*. From the perspective of the *Solanaceae* as a whole, the similar localization of apparent anthocyanin regulatory loci in tomato (*ag*), pepper (*A*), eggplant (*fap*10.1, *lla*10.1, *lra*10.1, *sa*10.1, *pa*10.1, *ca*10.1), and potato (*I*, *F*, *B*) within their respective genomes (i.e., close to RFLP marker TG233 or other markers known to be close to TG233) is fascinating. Once any of these genes has been isolated, it may be possible to rapidly isolate the remaining loci in this region through simple hybridization approaches.

Association between tomato *af* and *chi*

The recessive *anthocyanin free* (*af*) mutant of tomato has previously been shown to map to a 5 cM interval on chromosome 5, bounded to the north by CT93 and to the south by CT63A (Tanksley et al. 1992). In the current study, *chi* also mapped to this region, 0.7 cM south of CT93. The *af* phenotype is associated with the loss of both anthocyanin pigments and of trichomes and thus mirrors the well-known pleiotropic *transparent testa glabra 1* (*ttg1*) locus in *Arabidopsis* (Walker et al. 1999). The *ttg1* gene product is a WD40 repeat protein, very similar to the *Petunia* an11 gene product. Although *chi* and *af* map to the same region, it thus does not appear likely that *chi* will correspond to *af*.

Association between tomato *ae* and *ans*

The recessive *entirely anthocyaninless* (*ae*) mutant has previously been mapped to the lower arm of chromosome 8, approximately 4 cM south of CT265 and 12 cM north of CT68 (Tanksley et al. 1992). The *ans* gene maps to the same area, approximately 1 cM north of CT265. The *ans* gene thus constitutes a candidate for *ae*. The bHLH transcriptional regulator *jaf13* also maps to the same area, about 4.2 cM south of CT265 and 13.5 cM north of CT68 and constitutes a second candidate for *ae*. *Jaf13* is known to regulate expression of *dfr* in *Petunia* (Quattrocchio et al. 1988). Because *ae* does not abolish the enzymatic activity of DFR (Yoder et al. 1994), it seems unlikely that *jaf13* corresponds to *ae*. Transposon insertions in the maize *a2* gene, which also encodes *anthocyanidin synthase*, completely abolish anthocyanin accumulation (Menssen et al. 1990). Phenotypically, *ans* seems the better candidate for *ae*.

There are several classical tomato loci (e.g., *anthocyanin absent* [chromosome 2, lower arm], *baby lea syndrome* [3L], *anthocyanin loser* [8L], *Hoffman's anthocyaninless* [9L]) for which we were unable to identify compelling candidate genes. This resulted at least in part because the map locations of many classical loci (Mutschler et al. 1987) remain unmapped relative to RFLP markers. It is possible that *baby lea syndrome* corresponds to Petunia *an11*, that *anthocyanin loser* codes for a bHLH protein similar to *Petunia jaf13*, and that *Hoffman's anthocyaninless* corresponds to *Petunia rt* or *an1*, but since none of these three classical loci have been mapped relative to molecular markers, it is not possible to make compelling predictions with current mapping data.

Testing specific hypotheses by sequencing candidate genes

Excellent support for the hypotheses that *anthocyanin reduced* encodes F3H and that *anthocyaninless* encodes F3′5′H were obtained by virtue of having detected mutations in the *f3h* and *f3′5′h* genes, respectively, of mutant tomato accessions. A technical challenge we faced with *f3′5′h*, and that would also be faced in testing several other candidate gene hypotheses in tomato, is that tomato EST or gene sequences have not yet been described for all anthocyanin biosynthetic or regulatory genes (e.g., anthocyanidin synthase or a bHLH domain gene with similarity to *Petunia an1*). The complete absence of several tomato and potato anthocyanin genes in current sequence databases was rather unexpected since over 155,000 tomato and 90,000 potato ESTs have been deposited to date.

The large number of promising positional candidate genes identified in this work demonstrates that mapping genes known to influence a biochemical pathway can be a powerful means of identifying genes responsible for pathway-related natural variation. This mirrors the previous success of identifying candidates for important carotenoid accumulation loci by mapping all carotenoid biosynthetic structural genes in pepper (Thorup et al. 2000). In the present study, several regulatory loci were mapped in addition to many structural genes, which served to increase the number of interesting associations. The value of pathway mapping for both carotenoids and anthocyanins pathway was enhanced by the ability to interpret map data in other solanaceous crops through comparative genetic maps. As such maps expand to include additional species in the future, the added value from pathway mapping approaches should increase even further.

Acknowledgements We thank S.D. Tanksley for providing tomato mapping filters and cDNA clones. This work was supported in part by Federal Hatch funds provided to W.D.

References

Beld M, Martin C, Huits H, Stuitje AR, Gerats AGM (1989) Flavonoid synthesis in *Petunia hybrida*: partial characterization of dihydroflavonol-4-reductase genes. Plant Mol Biol 13:491–502

Ben Chaim A, Borovsky Y, De Jong W, Paran I. Linkage of the *A* locus for the presence of anthocyanin and *fs10.1*, a major fruit shape QTL in pepper. Theor Appl Genet 106:889–894

Bongue-Bartelsman M, O'Neill SD, Tong Y, Yoder JI (1994) Characterization of the gene encoding dihydroflavonol 4-reductase in tomato. Gene 138:153–157

Bonierbale M, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics 120:1095–1103

Borgnino F, Rick CM, Opena RT, Zobel RW (1973) Two more anthocyaninless loci for chromosome 2. Tomato Genetics Cooperative Report No. 23.

Daskalov S, Poulos JM (1994) Updated *Capsicum* gene list. Capsicum Eggplant Newsl 13:15–26

De Jong H (1991) Inheritance of anthocyanin pigmentation in the cultivated potato: a critical review. Am Pot J 68:585–593

De Jong WS, De Jong DM, De Jong H, Kalazich J, Bodis M (2003) An allele of dihydroflavonol 4-reductase associated with the ability to produce red anthocyanin pigments in potato (*Solanum tuberosum* L.). Theor Appl Genet (in press)

Dodds KS, Long DH (1955) The inheritance of colour in diploid potatoes. I. Types of anthocyanidins and their genetic loci. J Genet 53:136–149

Dodds KS, Long DH (1956) The inheritance of colour in diploid potatoes. II. Three-factor linkage group. J Genet 54:27–41

Doganlar S, Frary A, Daunay M-C., Lester RN, Tanksley SD (2002a) A comparative genetic linkage map of eggplant (*Solanum melongena*) and its implications for genome evolution in the *Solanaceae*. Genetics 161:1697–1711

Doganlar S, Frary A, Daunay M-C., Lester RN, Tanksley SD (2002b) Conservation of gene function in the *Solanaceae* as revealed by comparative mapping of domestication traits in eggplant. Genetics 161:1713–1726

Feinberg AP, Vogelstein B (1983) A technique for radiolabelling DNA restriction fragments to a high specific activity. Anal Biochem 132:6–13

Fulton TM, van der Hoeven R, Eannetta NT, Tanksley SD (2002) Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. Plant Cell 14:1457–1467

Gebhardt C, Ritter E, Barone A, Debener T, Walkmeier B, Schachtschabel U, Kaufmann H, Thompson RD, Bonierbale MW, Ganal MW, Tanklsey SD, Salamini F (1991) RFLP maps of potato and their alignment with the homeologous tomato genome. Theor Appl Genet 83:49–57

Goldsbrough A, Belzile F, Yoder JI (1994) Complementation of the tomato *anthocyanin without* (*aw*) mutant using the dihydroflavonol 4-reductase gene. Plant Physiol 105:491–496

Holton TA, Cornish EC (1995) Genetics and biochemistry of anthocyanin biosynthesis. Plant Cell 7:1071–1083

Holton TA, Brugliera F, Lester DR, Tanaka Y, Hyland CD, Menting JGT, Lu C-Y, Farcy E., Stevenson TW, Cornish EC (1993) Cloning and expression of cytochrome P450 genes controlling flower colour. Nature 366:276–279

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lewis CE, Walker JRL, Lancaster JE, Sutton KH (1998) Determination of anthocyanins, flavonoids and phenolic acids in potatoes. I. Coloured cultivars of *Solanum tuberosum* L. J Sci Food Agric 77:45–57

Livingstone KD, Lackney VK, Blauth JR, van Wijk R, Jahn MK (1999) Genome mapping in *Capsicum* and the evolution of genome structure in the *Solanaceae*. Genetics 152:1183–1202

Menssen A, Hohmann S, Martin W, Scanle PS, Peterson PA, Saedler H, Gierl A (1990) The En/Spm transposable element of *Zea mays* contains splice sites at the termini generating a novel intron from a dSpm element in the A2 gene. EMBO J 9:3051–3058

Mutschler MAM, Tanksley SD, Rick CM (1987) 1987 Linkage maps of the tomato (*Lycopersicon esculentum*). Tomato Genet Coop Rep 37

Noda Y, Kneyuki T, Igarashi K, Mori A, Packer L (2000) Antioxidant activity of nasunin, an anthocyanin in eggplant peels. Toxicology 148:119–123

O'Neill SD, Tong Y, Sporlein B, Forkmann G, Yoder JI (1990) Molecular genetic analysis of chalcone synthase in *Lycopersicon esculentum* and an anthocyanin-deficient mutant. Mol Gen Genet 224:279–288

Quattrocchio F, Wing JF, van der Woude K, Mol JNM, Koes R (1998) Analysis of bHLH and MYB domain proteins: species-specific regulatory differences are caused by divergent evolution of target anthocyanin genes. Plant J 13:475–488

Quattrocchio F, Wing JF, van der Woude K, Souer E, de Vetten N, Mol JNM, Koes R (1999) Molecular analysis of the *anthocyanin2* gene of petunia and its role in the evolution of flower color. Plant Cell 11:1433–1444

Selinger DA, Chandler VL (1999) Major recent and independent changes on levels and patterns of expression have occurred at the *b* gene, a regulatory locus in maize. Proc Nat Acad Sci USA 96:15007–15012

Spelt C, Quattrocchio F, Mol JNM, Koes R (2000) *anthocyanin1* of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. Plant Cell 12:1619–1631

Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovanonni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Roder MS, Wing RA, Wu W, Young ND (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Thorup TA, Tanyolac B, Livingstone KD, Popovsky S, Paran I, Jahn M (2000) Candidate gene analysis of organ pigmentation loci in the *Solanaceae*. Proc Natl Acad Sci USA 97:11192–11197

Van Eck HJ (1995) Localisation of morphological traits on the genetic map of potato using RFLP and isozyme markers. Ph.D thesis, Wageningen Agricultural University

Van Eck HJ, Jacobs JME, van Dijk J, Stiekema WJ, Jacobsen E (1993) Identification and mapping of three flower colour loci of potato (*S. tuberosum* L.) by RFLP analysis. Theor Appl Genet 86:295–300

Van Eck HJ, Jacobs JME, Van Den Berg PMMM, Stiekema WJ, Jacobsen E (1994) The inheritance of anthocyanin pigmentation in potato (*Solanum tuberosum* L.) and mapping of tuber skin colour loci using RFLPs. Heredity 73:410–421

Van Eck HJ, Rouppe van der Voort J, Draaistra J, van Zandvoort P, van Enckevort E, Segers B, Peleman J, Jacobsen E, Helder J, Bakker J (1995) The inheritance and chromosomal localisation of AFLP markers in a non-inbred potato offspring. Mol Breed 1:397–410

Vetten N de, Quattrocchio F, Mol J, Koes R (1997) The *an11* locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. Genes Dev 11:1422–1434

Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, Esch JJ, Marks AD, Gray JC (1999) The *transparent testa glabra1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes a WD40 repeat protein. Plant Cell 11:1337–1349

Wettstein-Knowles P von (1967) Mutations affecting anthocyanin synthesis in the tomato. I. Genetics, histology, and biochemistry. Heredity 60:317–346

Yoder JI, Belzile F, Tong Y, Goldsbrough A (1994) Visual markers for tomato derived from the anthocyanin biosynthetic pathway. Euphytica 79:163–167